数据库安全研究现状与展望 *

张 敏

（中国科学院软件研究所信息安全国家重点实验室 北京 100190）

摘要 数据库系统是当今大多数信息系统中数据存储和处理的核心。针对数据库系统的攻击将直接导致敏感与隐私信息泄露。开展数据库安全理论与技术研究，是实现数据库系统安全的迫切需求。文章通过总结与分析数据库安全的研究现状与发展趋势，认为我国在该领域与国际上的差距在不断缩小；当前既是数据库技术领域变革期，也是我国数据库安全领域研究的重要机遇期，并在此基础上提出了对策与建议。

关键词 数据库安全，安全数据库管理系统，数据库-即-服务，海量信息处理

DOI：10.3969/j.issn.1000-3045.2011.03.008

张敏副研究员

1 概述

数据库系统是当今大多数信息系统中数据存储和处理的核心。由于数据库中常常含有各类重要或敏感数据，如商业机密数据、个人隐私数据、甚至是涉及国家或军事秘密的重要数据等，且存储相对集中，因而针对数据库的攻击往往能达到最为直接的效果。例如，Verizon Business 的年度计算机破坏报告中提到，在近年的数据丢失案中，数据库破坏占据了 30%；而在数据入侵的统计中，数据库入侵则高达 75%。不仅如此，针对数据库系统的成功攻击往往导致黑客获得所在操作系统的管理权限，从而为整个信息系统带来更大程度的破坏，如服务器瘫痪、数据无法恢复等等。因此，及时开展数据库安全的理论与技术研究，实现数据库系统的安全防护，是数据库自出现以来就一直存在的迫切需求。

数据库安全研究的基本目标是研究如何利用信息安全及密码学技术，实现数据库内容的机密性、完整性与可用性保护，防止非授权的信息泄露、内容篡改以及拒绝服务。数据库安全是涉及信息安全技术领域与数据库技术领域的一个典型交叉学科，其发展历程与同时代的数据库技术、信息安全技术的发展趋势息息相关。在计算机单机时代、互联网时代以及当前的云计算时代，数据库安全需求发生了极大的变化，其内涵也更加丰富。本文将简要回顾数据库安全研究的发展历程与当前现状，并提出相关发展建议，希望对其未来发展起到一定参考与借鉴作用。

* 收稿日期：2011 年 5 月 11 日

2 安全数据库管理系统:计算机时代的数据库安全

早在上个世纪70年代，国际上数据库技术与计算机安全研究刚刚起步之时,数据库安全问题就引发了研究者的关注,相关研究几乎同步启动。当时的研究重点集中于设计安全的数据库管理系统,又称为多级安全数据库管理系统(Multi-Level Secure DBMS)。众所周知,数据库管理系统是负责数据存储、访问与管理的核心平台软件。因而它也理所当然成为维护数据库系统的安全核心。早期的数据库安全研究的核心目标在于,通过设计符合特定安全策略模型的安全数据库管理系统，严格实施访问控制策略、控制数据库内容的操作与访问,从而实现整个数据库系统的安全。相关研究大致可划分为如下几个重要阶段:

(1) 萌芽与初始时期:在上世纪70年代中期至80年代初,美国空军、海军资助的一批研究项目为多级安全数据库的研究奠定了基础。1975年,Hinke和Schaefer的报告给出了Hinke-Schaefer安全数据库研究的内容，实现了基于Multics操作系统的可信数据库管理系统;1983年,Woods Hole研讨班进一步提出了适用于多级安全数据库系统的三种体系结构：核心化(Hinke-Schaefer)结构、完整性锁结构和分布式结构。这个时期数据库安全研究的主流是军用安全数据库,美国军方的大力推动为研究工作涂上了一层浓重的军方背景。在萌芽与初始时期,研究者对数据库面临的安全威胁、数据库的安全需求以及安全数据库的研究问题有了基本的认识,通过若干项目的研发形成了安全数据库开发的方法论。

(2) 标准化时期:数据库安全研究进入标准化时期的重要标志是美国国防部计算机安全中心在1983年发表的可信计算机评估准则(TCSEC)。该准则于1985年被确定为美国国防部标准,是历史上第一个计算机安全评估准则。这一段时期是多级安全数据库系统发展的黄金时期,期间出现了一批达到高安全级别的数据库管理系统。其中比较有代表性的研究项目包括Seaview[1],ASD和LDV[2]。

Seaview(Secure Data View)是美国空军资助,SRI和Gemini公司共同参与的研究项目。其研究目标是实现一个达到TCSEC A1级的安全数据库，访问控制粒度达到字段级。Seaview采用了核心化的体系结构,由操作系统提供强制访问控制。ASD(Advanced Secure DBMS)与LDV(LOCK Data View)分别是美国TRW国防系统小组与美国空军资助进行的项目,安全性均达到了TCSEC标准A1级。此外,数据库软件开发商也积极响应,开发出一些安全等级稍低的安全数据库产品。如,Oracle的Trusted Oracle 7经评估达到B1级;Sybase的SQL Server达到了C2级,SQL Secure Server达到B1级且是最早通过B1级评估的安全数据库系统;Informix的INFORMIX-OnLine/Secure 5.0达到了B1级。

在标准化时期,围绕多级安全数据库管理系统的设计,形成了安全数据库的理论与技术基础,包括如下重要研究内容:① 数据库形式化安全数据模型分析及验证[3];② 数据库隐通道检测及其分析[4];③ 多级安全数据库事务模型及其分析[5];④ 数据库安全体系结构及实现技术;⑤ 数据库审计[6]与数据库加密等。

1991年,TCSEC关于数据库评估的解释(TNI)发表[7]。该文档详细说明了如何使用TCSEC标准对数据库管理系统和其他高级应用进行评估。它标志着研究者对于安全数据库的需求、功能、保证等达成了共识,一些关键技术进入了成熟阶段。

(3) 多样化时期：随着计算机网络技术的出现与发展，数据库所处的环境更为复杂，人们从实践中逐渐认识到，即使是严格按照数据模型设计实现的 MLS-DBMS，也并不能彻底解决数据库面临的各种安全威胁。同时数据安全研究中军方的色彩逐渐减退，而来自商业数据库的安全需求占据了主流，用户的安全需求也更为多样化。从此多级安全数据库步入了多样化时期。

具体来说，多样化体现在如下几个方面：①数据库应用环境和安全需求的多样化。基于军方安全需求的多级安全模型无法满足用户多样化、灵活的访问控制需求，出现基于角色的访问控制、基于属性的访问控制等细粒度访问控制，以及多策略访问控制框架等新型安全策略模型；②数据模型多样化。面向对象的数据库、XML 数据库等一批新型数据库系统的出现，打破了长期占据主流的关系数据模型，带动了半结构化数据访问控制模型研究；③信息安全技术体系多样化。随着信息保障(IA)概念的出现，以保护、检测、响应、恢复为核心内容的全生命周期的保护，取代了以往单一防护思想，引发了数据库入侵检测、可信恢复等相关研究内容。

在多样化时期，数据库安全研究内容更为广泛，典型的例子包括：①数据库细粒度访问控制模型研究[8]；②数据库入侵检测与恢复研究[9]；③数据库漏洞扫描技术研究；④SQL 注入攻击及防范技术研究等。

3 安全的数据库服务：互联网时代的数据库安全

互联网作为上世纪最伟大的发明之一，给社会带来了巨大的变化。在互联网时代，软件服务化逐渐发展成为一种为 IT 业界所广泛接受的工作模式。随着“软件-即-服务”理念的推广，越来越多的 IT 厂商选择将其非核心业务外包，从而集中更多的资源与精力投入到核心业务，达到降低成本、提高服务质量的目的。此外，近年来还出现了一批直接面对普通用户的数据库服务(或称“数据库-即-服务”，简称 DAS)，如亚马逊公司提供的 SimpleDB [10] 与 Relational Database Service (RDS)[11] 服务；谷歌公司推出的 Datastore [12] 服务；以及微软公司的 SQL Azure[13]服务等。这些数据库服务平台虽然采用不同的数据模型与实现技术，但都为用户提供快速、便捷的数据库服务，避免用户花费时间或精力用于软硬件采购与数据库日常维护管理。

一个典型的数据库服务场景由数据库内容提供者(简称所有者)、数据库服务运营服务商(简称服务者)与数据库使用者(用户)三方构成，如图 1 所示。

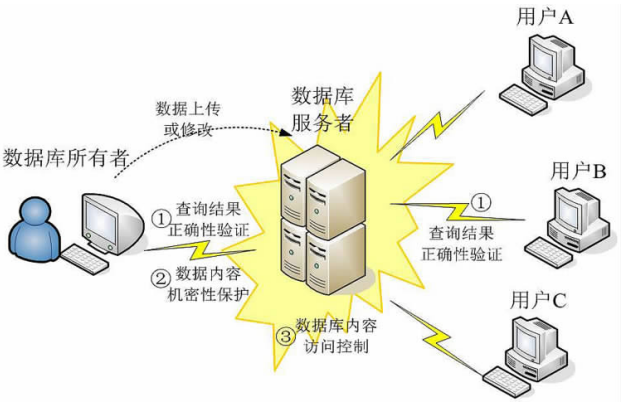

图 1 典型外包数据库场景及其安全需求

这种数据库服务模式带来了特殊的安全问题：数据库用户无法信赖安全数据库系统实施数据安全保护。因为在数据库服务模式下，由服务者负责维护数据库管理系统(DBMS)软件并提供数据库查询服务，但服务者并非完全可信，所以不仅外包数据库面临安全风险，DBMS 软件也因其运行的环境不可信、不可控而自身面临安全风险，无法

起到对数据的安全保护作用。这从根本上打破了以往的数据库安全威胁模型,带来了一系列安全问题。具体来说,“不可信的数据库服务者”这一安全假定引发了如下新问题:

(1)数据库所有者(DB Owner)的查询结果正确性验证需求。因为服务者不完全可信,所以用户与所有者在得到查询返回结果的同时,需要 DBMS 提供证据以表明结果的正确性。这包括两个方面:真实性与完备性。真实性即数据确实来自于数据库所有者,不是服务者伪造;完备性说明服务器返回了所有的正确内容,它没有为了提高系统吞吐率等原因只返回部分正确结果。

(2)数据库内容机密性保护需求。虽然数据库所有者委托第三方提供服务,但并不希望服务者知道所有数据内容,以防止其非授权泄漏并传播重要内容。因此要求服务者在部分内容是密文的前提下,仍然能够提供良好的数据库服务。

(3)来自所有者的数据库内容访问控制需求。数据库管理系统的一个基本安全需求是支持属主(所有者)对其内容进行授权与访问控制管理。在外包数据库系统中,数据库所有者需要技术证据证明第三方(服务者)正确执行了自己设定的访问控制策略。

(4)来自所有者的数据库内容版权证明需求。为了防止服务者的非法传播引发外包数据库所属权争议,所有者希望能在数据库中嵌入某些秘密,以向法庭证明自己对该数据库的所有权。目前数字水印技术是对多媒体数字作品进行版权保护的一种基本方法,但关系数据库是一种极为特殊的数据对象,是一个高度结构化的数据集合,与多媒体数字资源(图片、视频)相比存在很大差别,因而数据库水印与多媒体数据上的水印技术存在很大的不同。

目前,数据库服务模式下的数据库安全研究内容包括:①外包数据库安全检索技术研究[14];②外包数据库查询验证技术研究[15];③外包数据库密文访问控制技术研究[16];④数据库水印研究[17]等。

4 海量信息安全处理:云计算时代的数据库安全

当前,在 Web2.0 的背景下,互联网用户已由单纯的信息消费者变成了信息生产者,因而互联网上的信息呈爆炸式的速度增长。例如,Facebook 创始人兼首席执行官马克·扎克伯格(Mark Zuckerberg)在伦敦“互联网应用之未来”大会上指出,Facebook 用户共享个人信息需求的增长或存在着和“摩尔定律”类似的规律。美国国际数据公司一项名为“数字世界”的调查显示,2010 年全球共产生近 1.2 泽它(zetta,10 的 21 次方)字节的数字信息。而未来 10 年,全球总体信息量将是现在的 44 倍。毫无疑问,人类已经进入信息爆炸时代。在此背景下,支持海量数据高效存储与处理的云计算技术受到人们的广泛关注与青睐,在世界范围内得到迅猛发展,被誉为“信息技术领域正在发生的工业化革命”。

在云计算时代,信息的海量规模及快速增长为传统的数据库技术带来了巨大的冲击,主要挑战在于新的数据库应具备如下特性:① 支持快速读写、快速响应以提升用户的满意度;② 支撑 PB 级数据与百万级流量的海量信息处理能力;③ 具有高扩展性,易于大规模部署与管理;④ 成本低廉。在上述目标的驱使下,各类非关系型数据库(简称 NoSQL 数据库)应运而生,如:BigTable、HBase、Cassandra、SimpleDB、CouchDB、MongoDB 和 Redis 等。顾名思义,NoSQL 数据库为获得速度、可伸缩性及成本上的优势,放弃了关系数据库强大的 SQL 查询语言和事务机制。目前数据库领域关于未来

SQL 与 NoSQL 之中谁会消亡的讨论尚无定论，但近来 Twitter、Digg 和 Reddit 等多家 Web 2.0 企业宣布从 MySQL 转而使用 NoSQL 数据库，至少说明后者在现阶段更具商业潜力。

具体来说，在云计算时代，数据库安全研究面临如下新问题：

(1)海量信息安全检索需求。数据检索是用户最基本的需求之一，如关键词检索、全文检索、数据库 SQL 查询等。而海量数据的安全检索面临诸多挑战。一方面，现有的信息安全技术无法支持海量信息处理，例如数据经加密后丧失了许多原有特性，除非经过特殊设计，否则难以支持用户的各种检索；另一方面，当前的海量信息检索方法缺乏安全保护能力，例如当前的搜索引擎等不支持不同用户具有不同的检索权限。因此，如何在保证数据私密性的前提下，支持用户快速查询与搜索，是当前亟待解决的问题。

(2)海量信息存储验证需求。经典的签名算法与 HMAC 算法等均可用于验证某数据片段的完整性，但是当所需要验证的内容是海量信息时，上述验证方法需耗费大量的时间与带宽资源，以至于用户难以承受。因而在云计算环境下，数据库系统安全的需求之一是数据存在性与正确性的可信、高效的验证方法，能够以较少的带宽消耗和计算代价，通过某种知识证明协议或概率分析手段，以高置信概率判断远端数据是否存在并且未被破坏。

(3)海量数据隐私保护需求。海量信息带来的另一个重要问题就是隐私保护。与敏感信息不同，任何个体内容独立来看并不敏感，但是大量信息所代表的规律也属于用户隐私。例如，用户网上行为信息已成为一个潜力巨大的“金矿”。各大网站通过网络追踪技术记录用户的上网行为，分析用户偏好，并将上述信息高价出售给广告商，后者据此推送更精确的广告。因而在云计算环境下，研究如何抵抗从海量数据挖掘出隐私信息的方法，例如，将数据泛化、匿名化或加入适量噪声等等，对防止用户隐私信息泄露，具有重要的现实意义。

综上所述，在当前云计算模式下，数据库安全研究内容多集中在如下方面：①海量信息安全检索关键技术研究[18]；②海量数据完整性验证研究[19]；③海量数据隐私保护技术研究[20]。

5 国内研究现状与对策建议

5.1 国内研究现状

与国际同行相比，国内的数据库安全研究工作起步较晚，但也已取得了一定成绩。以安全数据库管理系统来说，目前已经达到国标 GB17859-1999 第三级（基本相当于 TCSEC B1 级）的国产数据库系统包括：可信 COBASE（北京大学、华中科大、人大）、达梦数据库系列（华中科大）、LOIS 安全数据库（中科院软件所信息安全国家重点实验室）、Softbase（南京大学）、Openbase Secure（东北大学）、神舟 OSCAR、以及 BeyonDB（中科院地理所、中科院软件所信息安全国家重点实验室）等。此外，在国家“863”计划的支持下，实现了国标第四级与第五级（基本相当于 TCSEC B2 级与 A1 级）安全数据库管理系统的关键技术研究与原型系统研发。总体来说，我国的科研人员已经掌握了数据库和数据库安全的关键技术，这为将来的研究开发工作奠定了基础。

5.2 建议与对策

目前我国在数据库安全领域的研究已经有了一定的基础，与国际差距正在逐渐缩小。但在基础理论研究、人才队伍建设以及成果转化等方面仍面临诸多困难。针对这些问题我们提出以下建议与对策：

(1)充分重视数据库安全所代表的交叉学科建设。数据库安全研究是涉及信息安全与数据库技术一个典型的交叉学科,需要两者之间的紧密结合。实际上,如未能充分理解数据库理论与实现技术,就无法设计出实用的数据库安全产品,如安全数据库管理系统。因此,从事数据库安全相关研究必须兼具两个领域的基本知识。而目前高校尚没有开设相关专业,信息安全专业学生普遍缺乏足够的数据库理论与技术基础。在人才培养过程中,需要花费大量时间与精力令其补充相关数据库知识。因此我们建议从学科建设角度出发,设立独立的数据库安全专业,组织出版相关的专业配套教材,这样可大大缩短人才培养时间,从而为未来培养大量数据库安全专门技术人才奠定基础。

(2)结合信息安全等级保护需求,推动国产数据库安全产品产业化。为切实提高我国信息安全保障水平,加强针对基础网络和重要信息系统的安全保护及管控能力,我国确立了信息安全等级保护制度,制订发布了一系列配套的标准规范,并开始在政府部门和各行业进行大规模的部署实施。这对于我国数据库安全技术研究以及国产安全数据库系统的开发与产品推广具有显著的推动作用,同时也是推动我国数据库安全产品自主化进程,提升自主安全产品和安全服务成熟度、适应性和集成能力的有效途径。

(3)加强先期项目投入与引导,把握重要历史机遇期赶超国际先进水平。由于历史原因,国外关系数据库产品占据了我国数据库市场的主要份额,这造成了我国的数据库安全技术发展缺乏足够的产业推动力。而当前云计算技术被誉为继互联网之后 IT 产业的第四次革新浪潮,为我国 IT 产业跨越升级发展提供了一个很好的机会。从国家安全战略角度出发,不应盲目追求引进国外的技术与产品,而应大力扶植发展具有我国自主知识产权的云数据安全管理技术,形成云计算数据安全国家标准,为实现我国云计算产业自主、可控提前抢占技术制高点。中科院作为我国科技领域的“国家队”,更应该充分利用前期技术积累,鼓励引导自主研发,培育云计算数据安全产学研联盟,抓住数据库技术升级换代的历史机遇,缩小该领域与国际同行之间的差距,实现跨越式发展。

主要参考文献

1 Lunt T F, Denning D E, Schell R R et al. The SeaView Security Model. IEEE Transactions of Software Engineering, 1990, 16(6): 593-607.

2 Stachour P D, Thuraisingham B. Design of LDV: A Multilevel Secure Relational Database Management System. IEEE Transactions on Knowledge and Data Engineering, 1990, 2(2): 190-209.

3 Jajodia S, Sandhu R S. Towards a Multilevel Secure Relational Model. SIGMOD, 1991, 20(2): 50-59.

4 McHugh J. Covert channel analysis: A chapter of the handbook for the computer security certification of trusted system. NRL Technical Memorandum. 1995, 5 540: 062A.

5 Jajodia S, Atluri V, Keefe T F et al. Multilevel Security Transaction Processing. Journal of Computer Security, 2001, 9(3): 165-195.

6 Agrawal R, Kiernan J, Srikant R et al. Hippocratic databases. In: proc. VLDB,2002.

7 National Computer Secuirty Center. Trusted Database Management System Interpretation of the TCSEC(TDI), NCSC-TG-021, 1991.

8 Rizvi S, Mendelzon A O, Sudarshan S et al. Extending Query Rewriting Techniques for Fine-grained Access Control. SIGMOD Conference, 2004, 551-562.

9 Liu P. Architectures for Intrusion Tolerant Database Systems. Proc. 18th Ann. Computer Security Applications Conf. (ACSAC 2002), 2002.

10 http://aws.amazon.com/simpledb/

11 http://aws.amazon.com/rds/

12 http://code.google.com/appengine/docs/java/datastore/

13 https://sql.azure.com/

14 Hacigumus H, Iyer B, Li C et al. Executing SQL over Encrypted Data in the Database Service Provider Model, SIGMOD, Madison, Wisconsin, USA, 2002.

15 Narasimha M, Tsudik G. DSAC: Integrity of outsourced databases with signature aggregation and chaining. In Proceedings of the ACM Conference on Information and Knowledge Management, 2005.

16 Vimercati S D C, Foresti S, Jajodia S et al. Over-encryption: Management of access control evolution on outsourced data. In Proc. of the 33rd VLDB Conference, Vienna, Austria, 2007.

17 Agrawal R, Kiernan J. Watermarking relational databases. In 28th Int'l Conference on Very LargeDatabases, Hong Kong, China, 2002.

18 Kamara S, Lauter K. Cryptographic cloud storage. Proceedings of the 14th international conference on Financial cryptograpy and data security. Tenerife. Canary Islands, Spain: Springer-Verlag, 2010, 136-149.

19 Bowers K D, Juels A, Oprea A. HAIL: a high-availability and integrity layer for cloud storage. Proceedings of the 16th ACM conference on Computer and communications security. Chicago, Illinois, USA: ACM, 2009, 187-198.

20 Chang C C, Thompson B, Wang H et al. Towards Publishing Recommendation Data With Predictive Anonymization, ASIACCS, 2010, 24-35.

Present Status and Trends of Researches on Database Security

Zhang Min

(State Key Laboratory of Information Security, Institute of Software, CAS 100190 Beijing)

Abstract Database system is the core part of data storage and processing in the current information systems. Attacks towards the database systems will directly result in the leakage of sensitive or private data. Research on the theory and techniques of database security is a basic requirement of database protection. By surveying on the present status and developments of the database security, we conclude that the gap between overseas and China is narrowing now, and the changing and developing of the database technology is providing a great opportunity for the database security researcher in China. Several suggestions and propositions are promoted to improve the researches in China.

Keywords database security, secure database management system, database-as-a-service, massive information processing

张 敏 中国科学院软件研究所信息安全国家重点实验室副研究员，博士，硕士生导师。主要从事数据库安全及数据隐私保护相关工作。发表论文 20 余篇，出版著作 1 部，获得国家发明专利授权 5 项。曾主持完成国家"863"计划重点项目课题、国防预研项目，参与完成"863"计划、国家发改委产业化项目、国家科技支撑项目等 10 余项。获得 2010 年度中国电子学会电子信息科学技术奖一等奖。主持研发的 BeyonDB 高可信地理空间数据库管理系统（安全版）与 LOIS 安全数据库管理系统达到国标 GB17859 第三级。E-mail: mzhang@is.iscas.ac.cn